

POLITECNICO DI TORINO
III Facoltà di Ingegneria
Corso di Laurea in Ingegneria Informatica

TESI DI LAUREA

Analisi delle reti neurali nello spazio
dei pesi e tecniche di riaddestramento

Relatore
prof. Elio Piccolo

Candidato
Bartolomeo Nicolotti

Settembre 2001

Introduzione

Nella tesi si è affrontato il seguente problema: come è possibile adattare una rete neurale al compito assegnato. Questo problema si presenta spesso nell'impiego di reti neurali, sia durante l'addestramento, non essendo noto a priori la difficoltà del compito, sia durante l'utilizzo della rete neurale, quando si incontrano nuove informazioni non coperte dal primo addestramento. Inoltre si è cercato di capire se al presentarsi di nuove informazioni convenga per adattare la rete ripartire da una nuova rete o meno.

Il problema può essere scomposto da un punto di vista teorico in due parti. La prima riguarda ciò che è possibile ottenere con le reti neurali, e qui i risultati sono incoraggianti in quanto si trova che con le reti neurali si possono approssimare compiti comunque complessi. Nulla in questa parte viene detto sulle dimensioni necessarie per ottenere l'approssimazione desiderata. La seconda parte riguarda la generalizzazione ottenibile in seguito ad un addestramento su campioni di esempio del compito. I risultati della teoria dell'apprendimento statistico sviluppata da Vapnik e Chervonensky ci dicono che per ottenere una buona generalizzazione occorre e basta che i parametri della rete siano in numero finito, nel senso che si esporrà in seguito. Bisogna quindi trovare un compromesso nello scegliere il numero di nodi per minimizzare l'errore nell'addestramento ed ottenere una buona generalizzazione.

Dal punto di vista pratico il problema più difficile da risolvere è quello di come addestrare la rete in modo efficiente. Per ottenere la rete neurale che svolga il compito richiesto occorre infatti stabilire i parametri della rete. Questo viene fatto nella fase di addestramento cercando di ridurre l'errore commesso dalla rete sui campioni a disposizione utilizzando algoritmi di ottimizzazione. Questi algoritmi non garantiscono di trovare il minimo globale della funzione d'errore se non per reti molto semplici. E' infatti garantito di trovare i parametri necessari solo per reti con un solo nodo, mentre la funzione di errore per reti di grandi dimensioni sembrano presentare molti minimi locali che rendono molto difficile ottenere il risultato cercato. Si è esaminato a fondo il più utilizzato metodo di addestramento, la propagazione all'indietro dell'errore (backpropagation) e le sue varianti per valutare le prestazioni e verificare i problemi di addestramento. Un altro aspetto critico dell'addestramento tramite backpropagation è l'estrema lentezza in prossimità di minimi locali e pianori e con un numero elevato di parametri da apprendere.

Per cercare la rete che sia un buon compromesso tra precisione sull'insieme di addestramento e buona generalizzazione si può cercare di adattare il numero di nodi della rete al problema. Un metodo per fare questo è quello di partire con un numero basso di nodi ed aggiungerne se dopo un certo numero di tentativi non si è raggiunto il risultato desiderato. Questo metodo è stato analizzato per capire se offre dei vantaggi rispetto ad altri.

Un metodo alternativo esaminato per l'addestramento è quello di suddividere i campioni in sottoinsiemi più semplici che possono essere appresi più facilmente. I sottoinsiemi vengono poi utilizzati per addestrare reti di dimensione più piccola che vengono composte per ottenere la rete cercata. Questo metodo è molto efficiente in termini di tempo, ma dal momento che la suddivisione viene fatta a caso si ottiene poca generalizzazione.

Si sono esaminate queste problematiche prima utilizzando compiti astratti, ma di semplice interpretazione, ed in seguito si è esaminata come semplice applicazione il riconoscimento di cifre manoscritte.